Predictive Analytics Using R and Python

- MUQUAYYAR AHMED
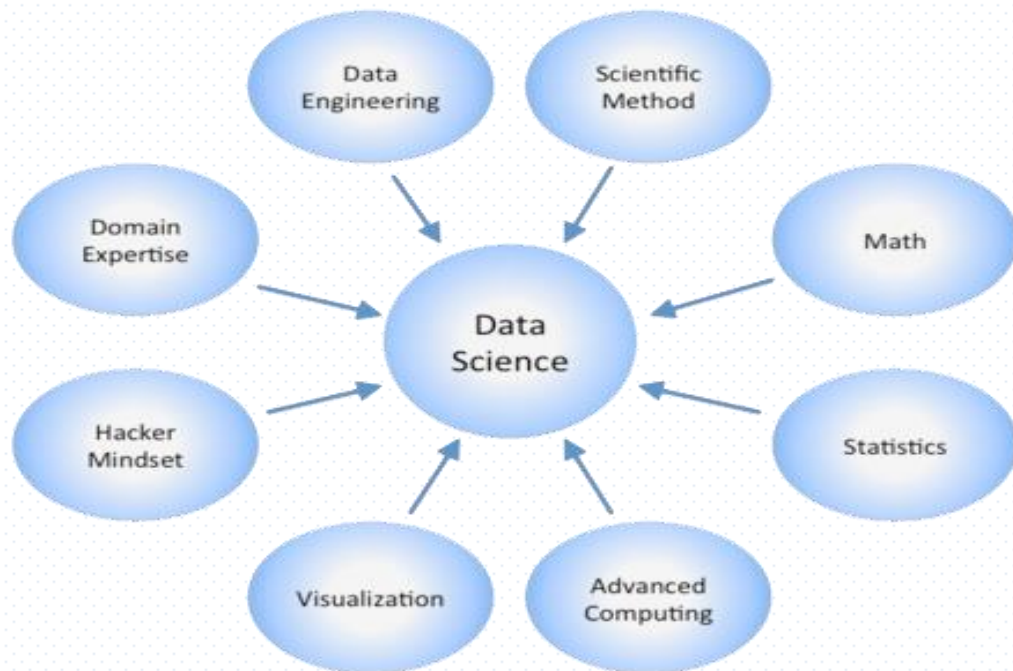 **DATA SCIENTIST**

Today's Agenda

Data Scientist!!!

Three skills
- Business understanding
- Technical
- Programming

Three tracks
- Engineering
- Services
- Consulting
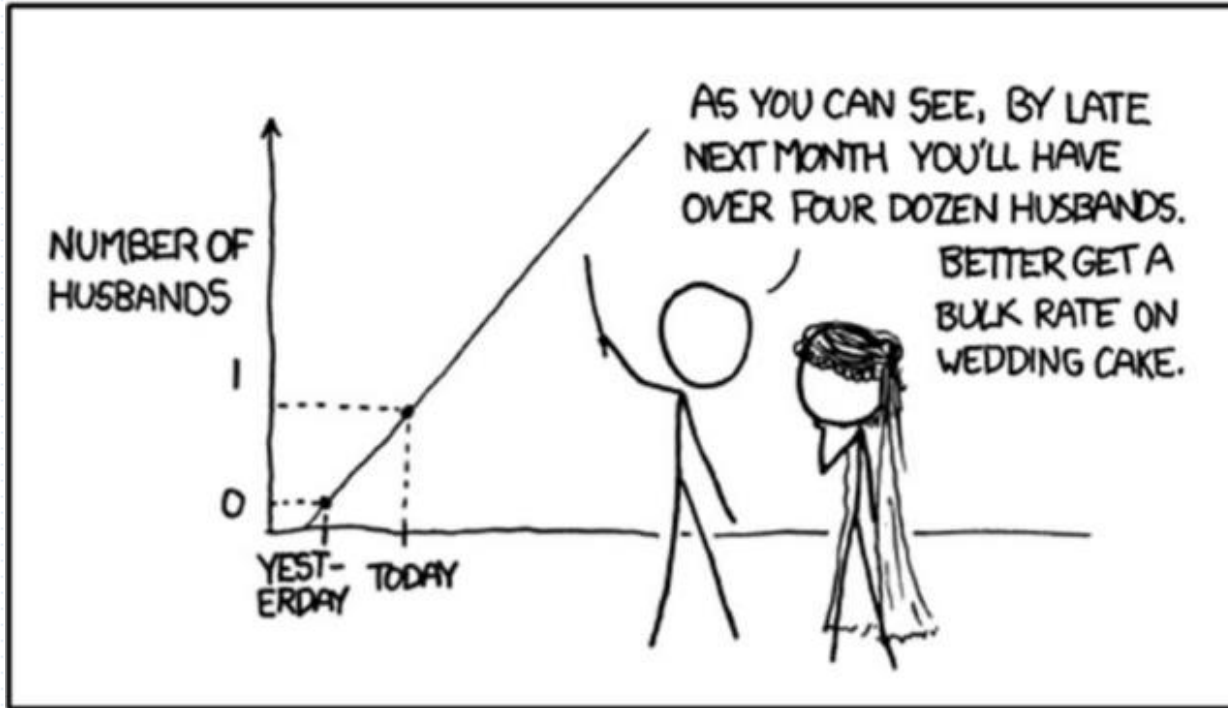
What is Data Analytics???

Data analytics is an emerging technique that dives into a data set without prior set of hypotheses

The data derive meaningful trends or hidden patterns that were not previously seen or empirically validated

Data analytics enables quick decisions or help change policies due to trends observed

Accumulation of raw data captured from various sources (i.e. discussion boards, emails, exam logs, chat logs in e-learning systems) can be used to identify fruitful patterns and relationships

Statistical modeling..

Type of problems Data Scientist deal with!!

Prediction/forecasting

Classification

Optimization

Unsupervised Learning

Cntd…

Problem Statement	Problem Category
A stock broker wants to figure out how much gold will cost tomorrow	Forecasting
A banker wants to find customers who are most likely to buy the new loan product he is launching	Classification
A hedge fund manager wants to figure out how to allocate the funds he has amongst the 10 different stocks he has chosen	Optimization
A chemist wants to find some interesting patterns in which patients are behaving upon administering the drug	Unsupervised Learning

Applications

Web search and information extraction

Banking and insurance

Pharmaceutical industry

Drug design

Predicting capital market behavior

Understanding customers

Designing robots and self-driving cars